UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.998a - 27

Research Workshop on Condensed Matter Physics
30 June - 22 August 1997
MINIWORKSHOP ON
QUANTUM MONTE CARLO SIMULATIONS OF LIQUIDS AND SOLIDS
30 JUNE - 11 JULY 1997
and
CONFERENCE ON
QUANTUM SOLIDS AND POLARIZED SYSTEMS
3 - 5 JULY 1997

"High-accuracy-Trotter-formula method for path integrals"

K. SCHMIDT
Arizona State University
Department of Physics & Astronomy
Tempe
AZ 85287-1504
U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

MAIN BUILDING STRADA COSTIERA, 11 TEL. 2240111 TELEFAX 224163 TELEX 460392 ADRIATICO GUEST HOUSE VIA GRIGNANO, 9 TEL.224241 TELEFAX 224531 TELEX 460449
MICROPROCESSOR LAB. VIA BEIRUT, 31 TEL. 2249911 TELEFAX 224600 TELEX 460392 GALILEO GUEST HOUSE VIA BEIRUT, 7 TEL.2240311 TELEFAX 2240310 TELEX 460392

High-accuracy Trotter-formula method for path integrals

K. E. Schmidt and Michael A. Lee*
Department of Physics and Astronomy, Arizona State University, Tempe, Arizona 85287
(Received 7 November 1994)

Path integrals are a powerful method for calculating real time, finite temperature, and ground state properties of quantum systems. By exploiting some remarkable properties of the symmetric Trotter formula and the discrete Fourier transform, we arrive at a high-accuracy method for removing "time slice" errors in Trotter-approximated propagators. We provide an explicit demonstration of the method applied to the two-body density matrix of ^{4}He. Our method is simultaneously fast, high precision, and computationally simple and can be applied to a wide variety of quantum propagators.

PACS number(s): 05.30.−d, 03.65.−w, 02.70.−c, 67.20.+k

Recent years have seen large scale computational methods that employ a "propagator" such as $\exp(-i\frac{Ht}{\hbar})$ or $\exp(-\frac{H}{k_BT})$ to carry out quantum time evolution [1,2] or calculations in quantum statistical mechanics [3,4]. This paper deals with analysis and enhancement of such propagator schemes when they are based on approximations like the Trotter formula [5],

$$e^{-\tau H} = \lim_{N\to\infty}\left[e^{-\tau T/N}e^{-\tau V/N}\right]^N. \tag{1}$$

(T is the kinetic energy and V the potential energy operator. The "time" variable τ is either $1/k_BT$ or $it/\hbar$.) We will show that a wide class of product approximations have errors proportional to only the *even* powers of $1/N$ and use this property with a Romberg-type integration scheme to successively eliminate time-slice errors to very high order. This approach permits substantial reduction of N, extending the variety of problems that are computationally accessible.

For numerical applications of Eq. (1) with finite N, errors depending on the time slice, $\Delta\tau = \tau/N$, enter through the break up,

$$e^{-\Delta\tau H} \approx e^{-\Delta\tau T}e^{-\Delta\tau V}, \tag{2}$$

which is in error at the level of $(\Delta\tau)^2$. The error accumulated in the N factors in the entire propagator is of order $1/N$. A superior break up was given by Feynman [6] who symmetrized the incremental propagator,

$$e^{-\Delta\tau H} \approx e^{-\frac{\Delta\tau}{2}V}e^{-\Delta\tau T}e^{-\frac{\Delta\tau}{2}V} \equiv U_S(\Delta\tau). \tag{3}$$

The error of the symmetrized propagator, U_S, is order $(\Delta\tau)^3$. This yields a $1/N^2$ error in a Green function or density matrix approximated by successive products of the form

$$\begin{aligned} e^{-\tau H} &\approx U_S(\tau/N)^N \\ &= e^{-\frac{\Delta\tau}{2}V}e^{-\Delta\tau T}e^{-\Delta\tau V}\cdots e^{-\Delta\tau V}e^{-\Delta\tau T}e^{-\frac{\Delta\tau}{2}V}. \end{aligned} \tag{4}$$

Many authors have offered error analyses similar to the above [4,6,7]. However, we found no attempt to exploit the error structure with standard extrapolation methods [8] to eliminate the low order errors resulting from discretization. For example, an operator without the $1/N^2$ error is

$$e^{-\tau H} \approx \left[4U_S(\tau/2N)^{2N} - U_S(\tau/N)^N\right]/3. \tag{5}$$

Implementing this numerically requires up to three times as much computation and double the storage, but the remaining error is of order $1/N^4$. As we show below, this process can be easily generalized so that each time the number of points is doubled, the error is again reduced by a factor of N^2. This enhanced error reduction occurs only when an asymptotic expansion for the error in the break up formula contains only powers of $1/N^2$.

Hatano and Suzuki [9] showed that the break up into N symmetric propagators, i.e., Eq. (4), does not include odd-order powers of $\Delta\tau$. Their proof involved an infinite expansion of commutators. We will show here that this error structure is more general. Our proof is based on the observation that if the propagator formed of N products is an even function of N its error expansion necessarily contains only powers of $1/N^2$. Given any break up, $e^{-\tau H} \approx U(\tau/N)^N$, evenness in N requires that it be invariant under $N \to -N$, i.e.,

$$[U(\tau/N)]^N = [U(-\tau/N)]^{-N}. \tag{6}$$

From this we see that an approximate propagator constructed of N products, each of which satisfies a "unitarity condition,"

$$U(\Delta\tau)U(-\Delta\tau) = 1, \tag{7}$$

will have an asymptotic error expansion with only powers of $1/N^2$.

The simple break up given by Eq. (2) fails to satisfy the unitarity condition. The symmetric propagator, U_S, does satisfy this as do several others, including the exact propagator, $e^{-\Delta\tau H}$, and certain rational approximations such as $(1-\Delta\tau H/2)/(1+\Delta\tau H/2)$. Indeed, if $U(\Delta\tau)$ does not satisfy the unitarity condition, then the construction $U(\Delta\tau/2)U^{-1}(-\Delta\tau/2)$ does.

*Permanent address: Department of Physics, Kent State University, Kent, OH 44242.

 ©1995 The American Physical Society

To illustrate the successive elimination of error terms in Trotter-approximated propagators, we apply our method to the calculation of the two-body density matrix of helium atoms and employ the symmetric break up of Eq. (3). The real-space two-body density matrix is a function of 12 coordinates. Eliminating the center of mass in the usual fashion leaves a function of six coordinates. The Bloch equation in these coordinates is then

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v(r)\right)\rho(\vec{r},\vec{r}',\beta) = -\frac{\partial}{\partial\beta}\rho(\vec{r},\vec{r}',\beta), \quad (8)$$

with the boundary condition $\rho(\vec{r},\vec{r}',\beta\to 0) = \delta^3(\vec{r}-\vec{r}')$. The Hamiltonian operates on the unprimed coordinates, m is the reduced mass and $v(r)$ is the central two-body potential which we take to be the HFDHE2 potential (a potential for He using the Hartree-Fock dispersion method) of Aziz *et al.* [10].

The usual partial wave decomposition can be used to expand in Legendre polynomials of the cosine of the angle between $\vec{r}$ and $\vec{r}'$,

$$\rho(\vec{r},\vec{r}',\beta) = \sum_{l=0}^{\infty}\frac{(2l+1)\rho_l(r,r',\beta)}{4\pi r r'}P_l(\hat{r}\cdot\hat{r}'). \quad (9)$$

The lth partial wave density matrix then satisfies a radial Bloch equation,

$$\left(-\frac{\hbar^2}{2m}\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2}\right] + v(r)\right)\rho_l(r,r',\beta) = -\frac{\partial}{\partial\beta}\rho_l(r,r',\beta) \quad (10)$$

with an analogous one-dimensional δ function boundary condition. When we refer to the "potential" for ρ_l below, we will implicitly include the r^{-2} term.

We have solved Eq. (10) for temperatures below 1 K, but we present results at 40 K because this temperature was selected in the bulk ^{4}He calculations of Ceperley and Pollock. They have discussed techniques employed to obtain the accurate two-body density matrix needed in their path integral Monte Carlo calculations [3] and adopt the "matrix squaring" method of Klemm and Storer [11]. We present numerical results for $l = 0$, but our method actually improves for larger l values and we routinely incorporate $l \geq 40$ in our construction of the full two-body density matrix.

Having specified our problem formally, we must develop a discrete representation for numerical solution. Following Eqs. (3) and (4), we break up the propagator for ρ_l into N equal slices of $\beta = 1/k_BT \equiv N\Delta\tau$ and choose a uniform spatial grid, so that the jth radial distance is $j\Delta r$, with $1 < j \leq M-1$, and $M\Delta r$ is large enough that the density matrix is numerically zero at that end point.

Evaluating any short-time propagator at the grid points produces a short-time density matrix,

$$U_{jm} = \langle j|U|m\rangle. \quad (11)$$

The discrete representation of the full density matrix becomes a series of matrix multiples,

$$\rho_l(j_N, j_0, \beta) = \sum_{j_1, j_2, \cdots, j_{N-1}} U_{j_N j_{N-1}}\cdots U_{j_2 j_1}U_{j_1 j_0}. \quad (12)$$

A small $\Delta\tau = 1/Nk_BT$ increment is equivalent to a large temperature so there are a variety of high temperature, semiclassical approximations that may be employed [3] to obtain the initial matrix in Eq. (11). However, simply evaluating a high temperature density matrix at the spatial grid points will not in general produce a matrix which satisfies the unitarity condition, Eq. (7), even if the original continuous density matrix $\langle r|U|r'\rangle$ does satisfy this condition.

Klemm and Storer [11] developed their "matrix squaring" method based on Eq. (12). If N is a power of 2, then Eq. (12) can be broken up into $\log_2 N$ factors, each one being the square of the previous one. Unless the unitary condition for the matrix U_{jm} is satisfied, the product approximation, Eq. (12), does not have the $1/N^2$ error structure.

A discrete space representation of the symmetric propagator that does satisfy the needed unitary condition can be calculated by Fourier transforms on a discrete basis. Examining Eq. (11) for the symmetric propagator, we see that the potential factors in Eq. (3) are diagonal in the real-space basis, but the kinetic energy operator is not. We insert a plane wave basis on M discrete k points $|k_n\rangle$ [with $k_n r_j = \pi nj/M$ and $\langle j|k_n\rangle = \sin(\pi nj/M)\sqrt{2/M}$]. The matrix element of a single symmetric propagator of Eq. (11) is then

$$U_{jm} = \sum_n \langle j|e^{-\frac{\Delta\tau}{2}V(r_j)}|k_n\rangle \times e^{-\Delta\tau\frac{\hbar^2k_n^2}{2m}}\langle k_n|e^{-\frac{\Delta\tau}{2}V(r_m)}|m\rangle. \quad (13)$$

This form regains the unitarity property, Eq. (7), and the products, Eq. (12), have even-order errors suitable for accurate extrapolation.

Using the unitary U_{jm}, the products for Eq. (12) can be carried out with direct matrix multiplication, the Klemm-Storer matrix squaring technique, or using fast Fourier transform (FFT) methods [2]. Our form of U_{jm} is well suited to the FFT method. This process is initiated with the infinite temperature density matrix $\rho_l(i,j,0) = \delta_{ij}/\Delta r$. Multiply it by $\exp\left[-\frac{\Delta\tau}{2}V(r_m)\right]$, then carry out a Fourier sine transform from space points r_m to points k_n. Multiply that result by the kinetic energy operator $\exp\left(-\Delta\tau\frac{\hbar^2k_n^2}{2m}\right)$, then Fourier transform back to the real-space points, r_j, and finally multiply by $\exp\left[-\frac{\Delta\tau}{2}V(r_j)\right]$. Repeat N times.

While our procedure for producing a discrete representation of the full propagators retains the unitarity property, it is nonetheless an approximation with error terms depending on the spatial grid size. We can show that this particular discretization is especially accurate because (perhaps surprisingly) each matrix product (including the FFT) is equivalent to using trapezoidal rule integration. The Euler-Maclaurin error formula for the

trapezoidal rule [12] states that the asymptotic expansion of the error is proportional to the odd derivatives of the function at the end points, plus a remainder term. All of our end point derivatives are zero to machine accuracy. Thus our grid errors rapidly become negligible once a sufficient number of integration points are employed, about 256 points in our case. Table I shows that spatial discretization errors do not undermine our assertions about the accuracy of extrapolating away time-slice errors.

We now detail the extrapolation in the discretization of "time" to achieve high accuracy and a resulting density matrix that is readily calculated to a precision of 10^{-9} using 64 bit arithmetic. The extrapolation formula of Eq. (5) illustrates the elimination of quadratic errors. Generalizing this is completely equivalent to performing the calculation for several different values of the discretization N, then fitting to a low order polynomial in $1/N^2$ to predict the exact result. We adopt a procedure exactly parallel to the Gear or Bulirsch-Stoer [8] methods for doing polynomial error estimates of integrating differential equations. This procedure involves the successive evaluation of the function for $N = 2, 4, 6, 8, 12, 16, 24, 32, \ldots$ points and a polynomial extrapolation is carried out until a prescribed error tolerance is met. We found it somewhat more efficient to break the full time interval into one to five equal segments, depending on the initial r' value, and in each segment apply the Gear prescription.

For illustrative purposes here, we drop the Gear procedure in order to graphically display the error versus time-step accuracy of the calculation. We start with N time slices, and repeat the calculation with $2N$, $4N$, $8N$, and $16N$ slices. Then Eq. (5) can be easily generalized to eliminate $1/N^2$, $1/N^4$, $1/N^6$, and $1/N^8$ errors. Figure 1 is a plot of the resulting errors in $\rho_0(r, r', \beta)$, for $r' = 3$ Å and $\beta = 0.025/\mathrm{K}$, with $N = 16$ initially. The logarithmic plot clearly illustrates that each removal of an additional power of $1/N^2$ increases accuracy by nearly two orders of magnitude. The ultimate accuracy possible is limited by the word length of the computer and is about half of the accuracy of the hardware. The top curve in the figure also illustrates that normally acceptable accuracy is achievable in calculations with discretizations as coarse as $N = 16$ and $N = 32$ with application of Eq. (5) to remove quadratic errors.

TABLE I. Comparison of the $l = 0$ radial density matrix $\rho_0(r, r', \beta)$ at $\beta = 0.025/\mathrm{K}$, for various numbers of integration points, M, and step sizes, Δr. The error tolerance for time extrapolations was set at 10^{-9}. All lengths are measured in angstroms. Our value of $\hbar^2/2m$ is 12.119 239 3 Å^2 K.

r	r'	M	Δr	$\rho_0(r, r', \beta)$
2.4	2.0	512	0.05	0.017 939 355
2.4	2.0	512	0.025	0.017 939 355
2.4	2.0	256	0.05	0.017 939 355
2.4	2.0	128	0.10	0.017 939 355
2.4	2.0	64	0.20	0.017 939 142
2.4	2.0	32	0.40	0.017 907 811
2.6	3.2	512	0.05	0.357 527 511
2.6	3.2	512	0.025	0.357 527 511
2.6	3.2	256	0.05	0.357 527 511
2.6	3.2	128	0.10	0.357 527 505
2.6	3.2	64	0.10	0.357 527 500
2.6	3.2	32	0.20	0.357 527 499

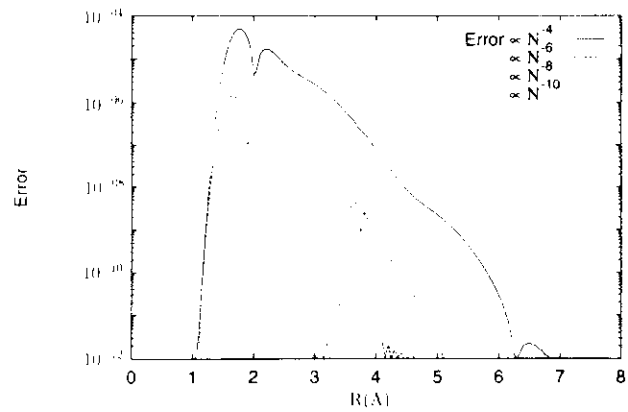

FIG. 1. Error after eliminating successive powers of $1/N^2$.

In Table I we provide a tabulation of ρ_0 to illustrate the error incurred due to k-space and r-space grid sizes. We see that the grid quickly becomes sufficiently dense that the error in the integration is essentially zero. The remarkable insensitivity of the result to the grid size follows from the Euler-Maclaurin error expansion for the trapezoidal rule.

We have performed the above calculations on other standard two-body He-He potentials, including the Lennard-Jones, hard sphere, and Morse potentials. All assertions in the above analysis remain true as long as the potential is sufficiently repulsive that the density matrix and its derivatives at the origin are small enough to be numerically ignored.

The use of the FFT is not essential to the extrapolation procedure, but because of the special form of the matrix U_{jm} in Eq. (13), the matrix products needed in Eq. (12) can be efficiently carried out. Matrix squaring is an alternative to the FFT, but the computational complexity of these approaches is different. Let N_S be the number of points in the spatial discretization and N_T be the number of terms in the break up formula. Our FFT method requires $N_T N_S^2 \ln N_S$ operations. Matrix squaring requires $N_S^3 \ln N_T$ operations. The N_S^2 in each of the above comes from needing $\rho(r, r', \beta)$ at $N_S \times N_S$ space points. This does not take into account the bandedness of the density matrix. For a given r', the only nonzero values of ρ are for positions r that are within a few thermal wavelengths of r'. (This range is easily estimated assuming free particle propagation.) If we label the number of nonzero values as N_B, then the above estimates change to $N_T N_S N_B \ln N_B$ for FFT, $N_S N_B^2 \ln N_T$ for matrix squaring. In any particular problem, it is likely that FFT will be successful, but a combination of FFT and matrix squaring followed by extrapolation will be optimal.

We have also investigated constructing the density matrix from an eigenfunction expansion, but found this im-

practical except when the eigenfunctions were known analytically. When numerical eigenfunctions are needed, the eigenfunction solver will (it is our experience) introduce unacceptable discretization errors. The complexity for evaluating an eigenfunction expansion requires $N_S^2 N_X$, where N_X is the number of terms needed in the expansion and $N_B N_S N_X$ when bandedness is exploited.

Our primary result is the demonstration that extrapolation to remove time-step errors is highly effective when the error structure of the propagator has only even-order terms. This result may be applied to the matrix squaring method as well. In that case, the cost of doubling the number of time slices is simply the cost of repeating the calculation with one more matrix multiply as compared with the FFT method where doubling the number of time slices actually doubles the cost. If the matrix squaring technique is based on a unitary propagator such as Eq. (3), and the propagator is discretized so that the matrix retains that unitarity, then extrapolation will be very cost effective. We would also point out that the matrix squaring technique can enjoy the same high order accuracy in the spatial grid discretization if the construction of the initial matrix is carried out using a procedure, for instance Eq. (13), where the trapezoidal rule integration scheme and Euler-Maclaurin error estimate applies.

In conclusion, we wish to reiterate that the even-order error structure of the propagator can be utilized by both real- and imaginary-time integration schemes. This allows many fewer time steps or greatly reduces the error for a moderate number of time steps. Our attention was originally drawn to this problem by De Raedt and Michielsen [1], who integrated the Schrödinger equation forward in real time to calculate quantum scattering. That work utilized a result of Suzuki [7] that combined forward and backward time integrations to make the error of a single composite time step fifth order, hence the total error proportional to $1/N^4$. We believe that our method could be used to produce higher-order real-time propagators. However, extrapolations like Eq. (5) produce a propagator that does not in general satisfy the unitarity condition. The violations of the unitary condition is, however, of the same order as the error. This means that our results may require the errors to be extrapolated to high precision to maintain stability for real-time propagations. This is similar to the Tchebychev method [2], which uses a polynomial approximation to the propagator $\exp(-H\Delta\tau)$. This method is also not unitary but the errors and the violation of unitarity are of the same order in $\Delta\tau$ as the order of the highest polynomial.

As our final observation, we note that the extrapolation procedure need not be applied directly to the calculation of an accurate propagator. Some techniques, such as Monte Carlo, do not employ a tabulated representation of the propagator. In such instances, the extrapolation scheme can be used much like standard finite-size scaling methods.

This work was supported by the NSF, Grant No. CHE-9407309. We acknowledge allocations from the Ohio Supercomputer Center (OSC), PGS178 and PGS030. The high precision work needed for the tables and figures was carried out on the OSC Cray YMP computer using double precision, i.e., 96 bit mantissa, arithmetic.

[1] H. D. Raedt and K. Michielsen, Comput. Phys. **8** (5), 600 (1994).
[2] C. Leforestier, R. H. Bisseling C. Cerjan, M. D. Feit, R. Friesner, A. Guldberg, A. Hammerich, and G. Jolicard, J. Comp. Phys. **94**, 59 (1991).
[3] D. M. Ceperley and E. L. Pollock, in *Monte Carlo Methods in Theoretical Physics*, edited by A. F. S. Caracciolo (ETS Editrice, Pisa, 1992).
[4] J. E. Hirsch, D. J. Scalapino, R. L. Sugar, and R. Blankenbecler, Phys. Rev. Lett. **47**, 1628 (1981).
[5] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[6] R. P. Feynman, Rev. Mod. Phys. **48**, 367 (1948).
[7] M. Suzuki, J. Math. Phys. **32**, 400 (1991).
[8] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1992).
[9] N. Hatano and M. Suzuki, Phys. Lett. A **153**, 191 (1991).
[10] R. Aziz, V. Nain, J. Carley, W. Taylor, and G. McConville, J. Chem. Phys. **70**, 4330 (1979).
[11] A. D. Klemm and R. G. Storer, Aust. J. Phys. **26**, 43 (1973).
[12] G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic Press, Orlando, 1985).

Green's Function Monte Carlo

in imaginary time. How to sample e^{-HT} with no time step error, — and what is the price?

The propagators

Schroedinger Equation in imaginary time:

$$(H - E_T)\Psi = -\frac{\partial \Psi}{\partial t}$$

$$\Psi(t+\Delta t) = e^{-(H-E_T)\Delta t}\,\Psi(t)$$

For small Δt, Euler's method would use

$$e^{-(H-E_T)\Delta t} \simeq 1 - (H-E_T)\Delta t$$

For continuous systems this is unstable since there are eigenstates with $E \to \infty$, and $1 - (H - E_T)\Delta t \longrightarrow -E\Delta t \longrightarrow -\infty$ for these states. If the system is on a lattice or otherwise has a maximum E, then if $1-(E-E_T)\Delta t > 0$ and ≤ 1 it is stable. If Δt is small enough, it works.

The backward or implicit Euler Equation would give the propagator

$$\frac{1}{1 + (H - E_T)\Delta t}$$

This is stable since a large positive E decays away.

The Kalos-Levesque-Verlet propagator

$$\frac{E_c + E_T}{E_c + H} = \frac{E_c + E_T}{E_c + E_T + H - E_T} = \frac{1}{1 + \frac{(H - E_T)}{E_c + E_T}}$$

E_c constant added to H

$$= \frac{1}{1 + (H - E_T)\Delta t} \quad \text{if} \quad E_c + E_T = \frac{1}{\Delta t}$$

Alternatively

$$= \frac{1}{\Delta t} \int_0^\infty dt \; e^{-(H - E_T)t} \; e^{-t/\Delta t}$$

↑

This is really what KLV method does, so we only need a slight change to make it the exact $e^{-(H - E_T)t}$ propagator.

All of the exact methods write the propagator as an approximate propagator plus a correction term.

The approximate propagator is chosen so that the correction term and itself are both positive. Monte Carlo sampling is used to select either the approximate propagator ($\Rightarrow$ a DMC like move), or the correction Term.

The exact propagator satisfies

$$(H - E_T)\, G(t) = -\frac{\partial G(t)}{\partial t} \quad \text{with } G(0) = 1.$$

I am not writing the space coordinates etc. Here for clarity.

An approximate propagator can be written as the solution of

$$\left(H^{(u)}(t) - E_T\right) G^{(u)}(t) = -\frac{\partial G^{(u)}(t)}{\partial t}$$

with $G^{(u)}(0) = 1$.

The identity

$$F(t) - F(0) = \int_0^t \frac{\partial}{\partial t'} F(t')\, dt'$$

becomes

$$G(t) = G^{(u)}(t) - \int_0^t dt' \frac{\partial}{\partial t'} \left[G(t-t')\, G^{(u)}(t')\right]$$

$$G(t) = G^{(u)}(t) + \int_0^t dt'\, G(t-t') \left[H^{(u)}(t') - H\right] G^{(u)}(t')$$

In DMC you want to pick $G^{(U)}$ so that the 2nd term is ≈ 0 and can be dropped. Typically you want the sign of the integrand of the second term to fluctuate so it is ≈ 0.

Here we want to apply Monte Carlo methods to both terms so we want both to be positive. Typically this means the effective time step is smaller for GFMC

But No extrapolation to $\Delta t \to 0$ is required.

Our equation corresponds to 2 processes. It says we can propagate according to $G^{(U)}(t)$, or we can propagate for a time $t' < t$ and scatter from $H^U - H$ and then propagate with the correct G for $t' >$ scattering time.

This is essentially Dyson's equation.

To make the notation easier, I will write these equations in a finite basis. You can think of solving for the ground state eigenvector/value of a large matrix, i.e. a lattice problem.

H_{ij} = matrix Hamiltonian

$G_{ij}(t)$ = matrix propagator as a function of t.

$G^{(0)}_{ij}(t)$ = approximate propagator.

The equation for G is

$$G_{ij}(t) = G^{(0)}_{ij}(t) + \int_0^t dt' \sum_{kl} G_{ik}(t-t') \cdot \left[H^{(0)}_{kl}(t') - H_{kl}\right] G^{(0)}_{lj}(t')$$

and for $\Psi(t)$

$$\Psi_i(t+\delta t) = \sum_j G_{ij}(\delta t)\, \Psi_j(t)$$

Importance Sampling

We will see how this helps us sample the G equation.

We need a trial function

$\Psi_T(i)$ = value of trial function in state i. (matrix index i)

Define

$$\tilde{\Psi}_m(t) = \Psi_T(m)\,\Psi_m(t)$$

$$\tilde{H}_{mn} = \frac{\Psi_T(m)\,H_{mn}}{\Psi_T(n)}$$

$$\tilde{G}^{(v)}_{mn}(t) = \frac{\Psi_T(m)\,G^{(v)}_{mn}(t)}{\Psi_T(n)}$$

$$\tilde{G}_{mn}(t) = \frac{\Psi_T(m)\,G_{mn}(t)}{\Psi_T(n)}$$

$$\tilde{G}_{ij}(t) = \tilde{G}^{(0)}_{ij}(t) + \int_0^t dt' \sum_{kl} \tilde{G}_{ik}(t-t') \cdot$$

$$\cdot \left[\tilde{H}^{(0)}_{kl}(t') - \tilde{H}_{kl} \right] \tilde{G}^{(0)}_{lj}(t')$$

$$\tilde{\Psi}_n(t+\Delta t) = \sum_m \tilde{G}_{nm}(\Delta t)\, \tilde{\Psi}_m(t)$$

Simplest Case — Assume $\tilde{G}$ is known.

We start with an ensemble of "walkers" Sampled from some distribution $\tilde{\Psi}_m(t)$

i.e. we have a set of values for m, the index. For each walker at m, we want to sample a new state n with probability proportional $p(n) = \dfrac{\tilde{G}_{nm}}{\sum_n \tilde{G}_{nm}}$. We must give this a "weight" $\sum_n \tilde{G}_{nm}$ on average. This is usually introduced using branching.

We want to choose $\Psi_T(m)$ to minimize this branching since this leads to higher variance ⟹ larger errors.

IDEAL IMPORTANCE SAMPLING

If $\Psi_T(m) = \Psi_0(m)$ → ground state

$E_T = E_0$ → ground state energy

$$\sum_n \tilde{G}_{nm}(t) = \sum_n \frac{\Psi_0(n)\, G_{nm}(t)}{\Psi_0(m)}$$

$$= 1$$

⟹ No branching, Zero variance of walker population.

Our G equation then must be the sum of 2 terms that add to 1.

We can therefore view the equation

$$\tilde{G}_{nm}(t) = \alpha\, p_1(n) + (1-\alpha)\, p_2(n)$$

- $\tilde{G}_{nm}(t)$: ↑ fixed t, fixed m
- $p_1(n)$: ↑ normalized prob.
- $p_2(n)$: ∫ term
- α: $G^{(u)}$ term; $0 \le \alpha \le 1$ if we have chosen $G^{(u)}$ properly.

The probability of picking the first term, α, should be

$$\alpha = \sum_n \tilde{G}^{(u)}_{nm}(t) \quad \text{and} \quad p_1(n) = \frac{\tilde{G}^{(u)}_{nm}(t)}{\sum_n \tilde{G}^{(u)}_{nm}(t)}$$

↑ Know everything about this term since it contains our known approximate $G^{(u)}$, and ψ_T.

The second term is

$$\int_0^t dt' \left[-\frac{\partial}{\partial t'} \sum_k \tilde{G}_{ik}(t-t') \, \tilde{G}^{(u)}_{kj}(t') \right]$$

The probability of selecting it is given by $1-\alpha$, and we can calculate that by first summing over i

$$\sum_i \tilde{G}_{ik} = 1 \quad \text{for ideal imp. sampling.}$$

The expected number of walkers is

$$\int_0^t dt' \left[-\frac{\partial}{\partial t'} \sum_k \tilde{G}^{(u)}_{kj}(t') \right]$$

and integrating gives $1-\alpha$. But notice this looks like the integral of a probability density in t'. We can view

$-\frac{\partial}{\partial t'} \sum_k \tilde{G}^{(u)}_{kj}(t')$ as a normalized probability density on $0 < t' < \infty$

If we pick $\tilde{G}^{(v)}(t \to \infty) = 0$.

Then $$\int_0^\infty dt' \underbrace{\left[\sum_k \tilde{G}^{(v)}_{kj}(t')\right]}_{p(t')} = 1$$

The probability that t' sampled from $p(t')$ on interval $0 < t' < \infty$ is more than t is

$$\int_t^\infty p(t')\,dt' = \alpha = \sum_k \tilde{G}^{(v)}_{kj}(t)$$

We can view $p(t')dt'$ as the probability that the walker scatters from the difference $H^{(v)}(t') - H$ at time between t' and $t'+dt'$.

The probability that the scattering takes place after t is just the probability of taking the first term.

For non ideal importance sampling,

$\sum_n \tilde{G}_{mn}(t) \neq 1$. We need to keep weights,

but we use $\sim$ to estimate the normalization

of the terms.

$$\tilde{G}_{ij}(t) = \left[\frac{\tilde{G}^{(u)}_{ij}(t)}{\sum_n \tilde{G}^{(u)}_{nj}(t)}\right]_2 \left[\sum_n \tilde{G}^{(u)}_{nj}(t)\right]_1$$

$$+ \int_0^t \sum_{kl} \left[\tilde{G}_{ik}(t-t')\right]_7 *$$

$$\left[\frac{\sum_{mn} (\tilde{H}^{(u)}_{mn}(t') - \tilde{H}_{mn})\, \tilde{G}^{(u)}_{nj}(t')}{\sum_{mn} (\tilde{H}^{(u)}_{mn}(t') - E_T \delta_{mn})\, \tilde{G}^{(u)}_{nj}(t')}\right]_6$$

$$\left[\frac{\tilde{H}^{(u)}_{kl}(t') - \tilde{H}_{kl}}{\sum_m (\tilde{H}^{(u)}_{ml} - \tilde{H}_{ml})}\right]_5 \left[\frac{\sum_m (\tilde{H}^{(u)}_{ml}(t') - \tilde{H}_{ml})\, \tilde{G}^{(u)}_{lj}(t')}{\sum_{mn} (\tilde{H}^{(u)}_{mn}(t') - \tilde{H}_{mn})\, \tilde{G}^{(u)}_{mn}(t')}\right]_4$$

$$\left[-\frac{\partial}{\partial t'} \sum_k \tilde{G}^{(u)}_{nj}(t')\, dt'\right]_3$$

Term 1 cancels denominator of term 2

Term 3 cancels " " term 6

H part of numerator of term 4 cancels the denominator of term 5

Denominator of term 4 cancels numerator of 6.

The Algorithm for exact sampling:

1. For each walker at position j that needs to be propagated, the time left is t. (i.e. initially a full time step)

2. Sample a time t' from the p.d.f. of term 3 $(0 < t' < \infty)$.
   The probability that $t' > t$ is term 1.
   If $t' > t$ sample a new state i from term 2. The walker has propagated to time t, and its weight is 1. It is completed for this time step.

3. If $t' < t$, then we have sampled a time t' from term 3. We sample the intermediate state l from the normalized prob. distribution of term 4, using the sampled t' value.

4. Given l, sample k from term 5. We have taken care of all terms except 6 and 7. Term 6 is now a number, and is taken to be the weight of the walker.

If perfect importance sampling is used this weight = 1, since $E_n^{local} \equiv \sum_m \tilde{H}_{mn}$ is equal to the ground state energy E_0.

For realistic Ψ_T, the weight fluctuates $\Rightarrow$ Branching as in DMC.

5. Only term 7 is left. We have a walker (s) in state k, and term 7 says it needs to be propagated from k to some final state i in time $t-t'$.

   To sample term 7, we simply reset the time left t to $t-t'$ and repeat steps 2 on. Eventually $t' > t$, and the walk terminates.

There are other breakups of the Equation that have the same limit when the importance sampling is ideal.

The KLV method is given in

Kalos, Levesque, Verlet Phys. Rev. A9 2178 (1974)

More details are in

M.E.S. Lecture Notes in Physics Vol 273
(Springer, 1987).

To give a concrete example, Let me describe the 2-d Heisenberg Spin 1/2 model on a square lattice. (Apologies to organizers)

$$H = J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j$$

↑ nearest neighbors.

This maps directly (with a small gauge transformation) onto a Bose problem. If you take ↑ spins to be particles, ↓ spins to be no particle, then you have a maximum of 1 particle/site ⇒ hard core.

The number of bosons ⇒ S_z. $S_z = 0$ = ½ filled.

$$H = -\frac{J}{2}\sum_{\langle ij\rangle}\left(b_i^+ b_j + b_j^+ b_i\right) \leftarrow \text{hopping}$$

$$+ J\sum_{\langle ij\rangle} n_i n_j + E_c$$

$\uparrow$ nearest neighbor repulsion. $\qquad \uparrow$ constant

A very simple trial function is

$$\psi_T(m) = e^{-\alpha N(m)}$$

where $N(m)$ = number of nearest neighbor pairs in state m.

The simplest possible $G^{(0)}$ is just take $H^{(0)}$ = diagonal part of H.

$$G_{ij}^{(0)}(t) = \exp\left[-(H_{jj} - E_T)t\right]\delta_{ij}$$

Since $G^{(0)}$ is diagonal

$$\tilde{G}_{ij}^{(0)}(t) = G_{ij}^{(0)}(t)$$

Terms 4 and 2 are just Kronecker deltas, and the equation becomes

$$\tilde{G}_{ij}(t) = \delta_{ij}\left[\exp(-H_{jj}-E_T)t\right]_1$$
$$+ \int_0^t dt' \sum_k \left[\tilde{G}_{ik}(t-t')\right]_7 \left[\frac{H_{jj}-E_j^{local}}{H_{jj}-E_T}\right]_6$$
$$\cdot \left[\frac{-\tilde{H}_{kj}(1-\delta_{kj})}{\sum_m(-\tilde{H}_{mj})(1-\delta_{mj})}\right]_5 \left[(H_{jj}-E_T)\, e^{-(H_{jj}-E_T)t'}\right]_3$$

Algorithm is

1. Sample a time t' from term 3. This is the time to hop. If $t' > t$, no particle hops, and we simply leave the walker at its current position and it has completed its walk.

2. If $t' < t$, then a particle hops. term 5 is the relative probability of a particle hopping in each possible direction. (If $\psi_T = 1$, all allowed hops would be equally probable). Sample a hop from term 5.

3. The weight is term 6. If $\Psi_T = \Psi_0$ and $E_T = E_0$ then this will be 1 as required for ideal importance sampling. Branch etc. as required and reset time to $t - t'$ and repeat steps 1 →.

Conclusion

I have given the general method for sampling $e^{-(H-E_T)t}$ with no time step error.

Advantages are no $\Delta t \to 0$ extrapolation required.

Disadvantage, slightly more complex $G^{(\omega)}$ are needed to guarantee all terms are positive. This may make algorithm slower than methods that have Δt error < statistical error. It also makes it a little more difficult to implement

($\approx$ 100 more lines in a GFMC versus DMC code).